THE JOURNAL OF NAVIGATION (2018), **71**, 419–436.
doi:10.1017/S0373463317000601

Analysis of Ship Accidents in the Istanbul Strait Using Neuro-Fuzzy and Genetically Optimised Fuzzy Classifiers

Sercan Erol[1], Mengü Demir[2], Bayram Çetişli[3] and Ekrem Eyüboğlu[1]

[1](*Department of Maritime Transportation and Management, Karadeniz Technical University, TR*)

[2](*Information Technology Department Directorate, Karadeniz Technical University, TR*)

[3](*Electrical Electronics Engineer, Isparta, TR*)

(E-mail: sercerol@ktu.edu.tr)

Marine accident analysis is important for ships passing through narrow, shallow and busy waterways. This study analyses the accidents which have occurred in the Istanbul Strait and proposes both quantitative and qualitative assessments of marine accidents. Marine accidents occurring in the Istanbul Strait are analysed by using a method based on neuro-fuzzy and genetically optimised fuzzy classifiers. It can be concluded that accident severity increases when poor weather conditions prevail in the Strait regardless of ship size. Therefore, solutions to reduce unwanted events should be prioritised by accounting for weather conditions and the capacity of the vessels. This analysis indicates that the safety level would be significantly improved if all the vessels follow the passage guidelines.

KEYWORDS

1. Accident Analysis. 2. Istanbul Strait. 3. Adaptive Neuro-Fuzzy Classifier. 4. Maritime Transportation. 5. Genetically Optimized Fuzzy Classifier.

Submitted: 7 June 2016. Accepted: 31 July 2017. First published online: 14 September 2017.

1. INTRODUCTION. The phenomenon of globalisation, together with developments in economics and technology, has had a great influence on many fields including world trade. This has caused an increase in world trade volume and correspondingly, an increase in marine transportation volumes. The Istanbul Strait has a strategic importance for the Caucasian-Caspian basin and countries having a Black Sea coastline. They bind the Aegean, Marmara and Black Sea together and for the aspect of marine transportation, have geopolitical importance and traffic density that makes them some of the most important sea corridors in the world.

In contrast to the developments in technology and enacted international safety regulations, marine accidents in the Istanbul Strait still pose a serious problem (Çelik and

Çebi, 2009). Studies show that, despite precautions, some accidents could not be prevented (Portela, 2005). In published research, it is stated that 70%–80% of marine accidents arise from human error (Harrald et al., 1998; Reason, 1997; Wang et al., 2005; Portela, 2005; Barnett, 2005; Trucco et al., 2007; Martins and Maturana, 2010; Başar and Yıldırım, 2014; Yıldırım et al., 2015; Erol and Başar, 2015; Uğurlu et al., 2015). Coastal areas, especially passages such as the Istanbul Strait are considered to be the most dangerous marine spaces (Uğurlu et al., 2013). Marine accidents occurring in coastal areas do great harm to humans, environment and economy (Uğurlu et al., 2015; Wei et al., 2015; Goerlandt et al., 2015). Indeed, many parties are directly affected by marine accidents such as the flag state, ship crew, other ships, environment, fishermen, cargo owners, ship owners, class societies, ship construction industry and insurance companies (Poyraz, 1998; Erol and Başar, 2015). For the Istanbul Strait, it is necessary to minimise the number of accidents to improve safety and limit damage and cost (Başar, 2010).

This paper aims to analyse maritime accidents and to improve safety and reduce the possibility of an unwanted event in the Istanbul Strait. Additionally, results obtained from this study can be an important input for risk assessment. Also, in a risk/ship accident analysis, risk is described from various perspectives (Goerlandt and Montewka, 2015). In this study the term risk is used as the possibility of an unwanted event. Accident archives of the Turkish Republic's Ministry of Transport, Maritime affairs and Communications and weather condition data from the Turkish State Meteorological Service have been used to prepare the model input and the International Maritime Organization (IMO) definition of accidents has been used to prepare the model output.

The remainder of this paper is organised as follows. The subject and method of study is mentioned together with previously conducted studies in Section 2. In Section 3, brief information is given about the Istanbul Strait. The data set and methods used are defined in Section 4. In Section 5, findings and results are discussed considering the rules generated for marine accidents on the Istanbul Strait. Furthermore, some datasets tested with Artificial Intelligence (AI) methods are summarised in this section. Finally, we conclude by giving a summary of the implementation in Section 5.

2. RELATED WORK.

Identifying accident factors in relation to marine accidents may contribute to the solution for lowering the count of accidents that affect many parties. AI methods, such as Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Classifier (ANFC), and Genetically Optimised Fuzzy Classifier (GOFC) can be used for analysing marine accidents.

There are some limited studies on marine accidents analysis by AI. Some recent works include the following. Hashemi et al. (1995) used a neural network for transport safety modelling. Le Blanc et al. (2001) carried out accident analysis of over 900 marine accidents that occurred on the lower Mississippi River, using neural computing techniques. Çelik and Çebi (2009) used a human factor analysis and classification system, based on a fuzzy analytical hierarchy process, to identify the role of human errors in shipping accidents. Qu et al. (2011) examined fuzzy ship domains based on a vessel collision risk assessment for the Singapore Strait. Küçükosmanoğlu (2012) developed an ANN model and Adaptive Neuro-Fuzzy Inference System (ANFIS) to forecast maritime accidents in the Istanbul Strait. This work defined the critical zones and the collision/capsizing accidents at high intensity zones in the strait. Mentes et al. (2015) identified and evaluated driving factors such as geographical locations at the time of the incident and failure modes causing fatalities for cargo ships using fuzzy set theory in coastal and open sea areas of Turkey.

Cetisli (2010a) proposed ANFC, using Linguistic Hedges (LH) which are fuzzy rule-based methods. These methods are used to classify the Istanbul Strait data. LHs improve the fuzzy meaning. The obtained fuzzy rule knowledgebase can easily be interpreted and shows the relation between input and output variables. To determine a proper maritime accident type, there are not enough data for some accident types such as machinery damage. Therefore, the number of samples in some accident types is very low. This situation affects the classification success. To eliminate this problem, the "Give&Take" validation method was also used. The Give&Take method creates two similar datasets from all the data (Cetisli and Kalkan, 2014). The effects of the features of the accident type determination were also analysed in this study.

There are also studies on marine accidents analysis on the Istanbul Strait, but the early work is somewhat limited. Akten (2004) reviewed 461 marine accident reports for accidents that occurred in the Istanbul Strait between the years 1953 and 2002. He organised the accidents by accident type and identified those sea areas that had the highest accident incident rates. Kum et al. (2008) investigated the risk profile of maritime accidents in the Istanbul Strait and then developed a methodology to minimise human error. Arslan and Turan (2009) examined the factors that influenced the occurrence of marine accidents in the Istanbul Strait. Ulusçu et al. (2009) analysed safety risks pertaining to transit vessel maritime traffic in the Istanbul Strait and proposed ways to mitigate marine accidents. Yazici and Otay (2009) developed a real time maritime traffic support model for safe navigation in the Istanbul Strait using a new MATLAB code. They concluded that local traffic density and pilotage turned out to be two main factors affecting the risks in the Istanbul Strait. Aydogdu et al. (2010) designated the most hazardous areas and identified the dangers regarding vessel traffic in the southern entrance to the Istanbul Strait. Erol and Başar (2015) analysed safety of life and economic loss in marine accidents occurring in the Turkish Straits using decision trees. Uğurlu et al. (2015) focused on marine accidents in the Turkish Straits that did serious harm to humans, the natural environment, and the economy.

3. ISTANBUL STRAIT.

Istanbul is the only city in the world that stands astride two continents and is without any alternative waterway. Therefore, the Istanbul Strait, the narrow waterway separating Europe from Asia, holds a strategic importance in maritime transportation as it links the Black Sea to the Mediterranean and the open ocean beyond. It is a vital passageway not just for trade but for the projection of military and political power (Ulusçu et al., 2009; Ucan and Nas, 2016).

The Istanbul Strait is distinct among the waterways of the world in its geographical structure and oceanographic characteristics. It is an approximately 30 km long, geographically narrow sea corridor which has sharp turns. Its mean width is 1,500 m and its narrowest place is the Kandilli region at 696 m. In the shallowest area, depth is 19 m. There are a total of 12 sharp turns throughout the strait and the sharpest turn is Yeniköy at 80^0. Deep and surface currents that rise to 6–8 knots from time to time significantly affect navigation safety. Traffic density of the Istanbul Strait is presented in Figure 1.

There is a two layered current pattern in the strait. The current from the Black Sea to the Marmara Sea flows on the surface and Mediterranean water flows from the Marmara

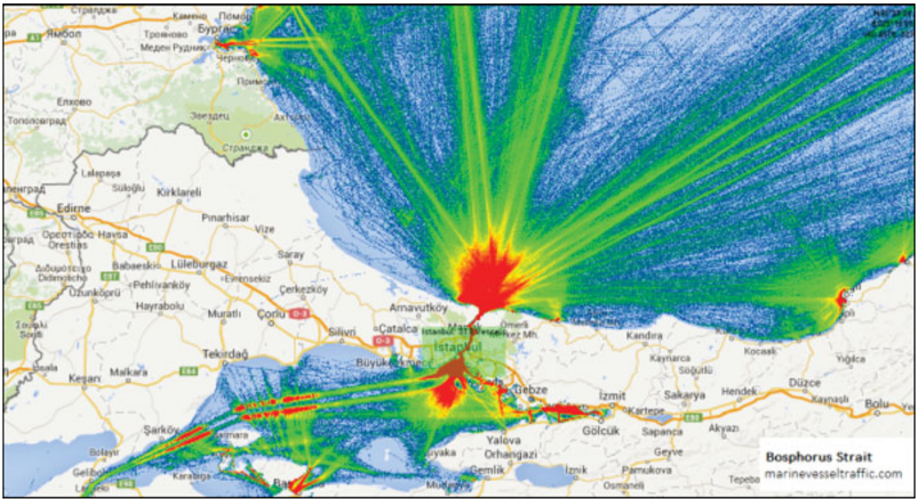

Figure 1. Traffic density of Istanbul Strait (Marine Traffic, 2016).

Sea to the Black Sea at the bottom of the strait. Sometimes, the reverse currents can also be seen on the surface (Küçükosmanoğlu, 2012). In addition to the geographical, physical and oceanographic structure of the strait; traffic density, dangerous cargo transportation, increasing ship lengths, complicated traffic system, sensitive environmental conditions, local dangers, the other activities affecting ship traffic, rising sea accidents, port configuration and bridges are also issues that have an influence on safe navigation throughout the Istanbul Strait (Uğurlu et al., 2015). Thus, the Istanbul Strait is considered to be one of the world's most dangerous waterways to navigate. Over 50,000 transit vessels pass through the Strait annually, 20% of which carry dangerous cargo (Ulusçu et al., 2008).

The legal framework of the transit traffic through the Strait of Istanbul is governed by the 1936 Montreux Convention. The vessels in question shall not be required to make any other stop during their passage through the Strait and pilotage and towage remains optional (Montreux Convention, 1936. Article 3–4). Accidents during oil transportation in the Strait, in particular, threaten serious ecological problems and impacts on people in Istanbul. Due to these problems, Turkey introduced traffic separation schemes in 1994, which were reorganised in 1998 and named Maritime Traffic Regulations for the Turkish Straits (MTRTS), to enhance safety of navigation (MTRTS, 1998).

4. MATERIAL AND METHODS. This section first defines the data set used in the study and then explains the methods which depend on artificial intelligence.

4.1. *Data Set.* In this study, maritime accident data has been taken from the accident archives of the Ministry of Transport, Maritime and Communications Search and Rescue Department of the Turkish Republic (MTMC-SRD). The accident data includes information about ships' names, flags, types and tonnage, coordinates of accidents, reasons for accidents, results of accidents, size of accidents, time of accidents (if available) and impact of accidents to the crew and economy. Weather condition data has been taken from the Kireçburnu, Kumköy and Göztepe stations of the Turkish State Meteorological Service (TSMS). All data are registered and official. The basic parameters used for the analysis are also Gross Registered Tonnage (GRT), ship Length Overall (LOA), Sea Wave Height

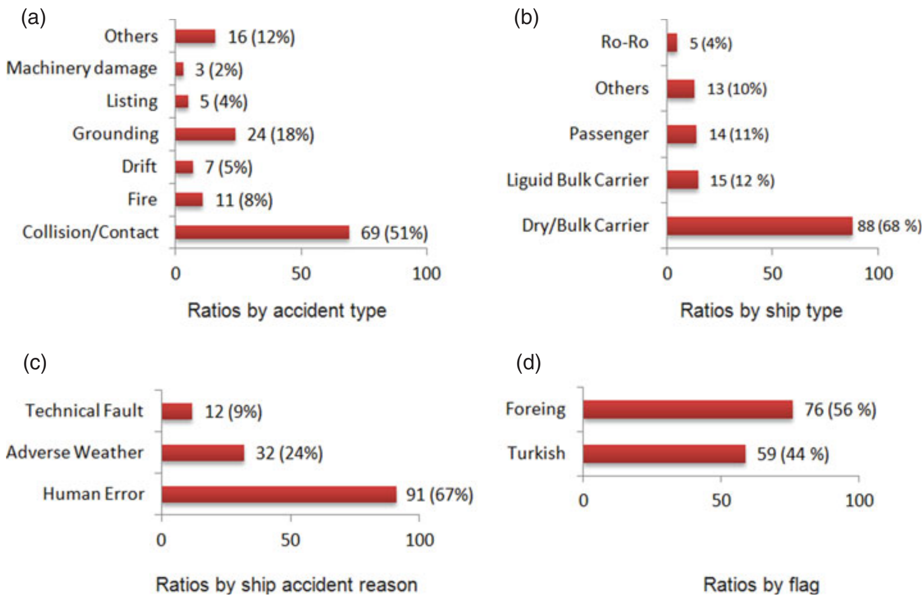

Figure 2. Descriptive statistics of the marine accident data.

Table 1. Description of height of accidents according to TSMS.

Sea State Code	Characteristics	Height (m)
0	calm (glassy)	0
1	calm (choppy)	0–0·1
2	flat (ripple)	0·1–0·5
3	light	0·5–1·25
4	temperate	81·25–2·5
5	rough sea	2·5–4
6	very rough sea	4–6
7	high rough sea	6–9
8	very high rough sea	9–14
9	terrible	>14

(LW) and Daily Average Wind Speed (DAWS). All of them are quantitative. According to the data set, although 880 marine accidents were recorded between 2001 and 2015, only 135 marine accidents matched with weather conditions data are used. This is because the majority of these accidents are very minor marine incidents.

Marine accident data has four ratings for accidents which are 'Less Serious', 'Serious', 'Very Serious' and 'Marine Incident'. Marine accident data used in the chart are grouped according to accident type, accident reason, ship type, and ship flag. The descriptive statistics of the data are given in Figure 2.

According to the Turkish State Meteorological Service, winds that have a velocity between 10·8 m/s and 17·1 m/s are considered 'strong winds' and winds that have a velocity of more than 17·2 m/s are considered 'gale'. The descriptive lengths of waves are given in Table 1.

The analysed marine accidents have been classified by the IMO's identification indicated in Table 2.

Table 2. Severities of accidents according to IMO (MSC-MEPC.3/Circ. 1).

Casualty classes	
Very serious accidents (VSA)	Total loss of the ship, loss of life, or severe pollution.
Serious accidents (SA)	Involving a fire, explosion, collision, grounding, contact, heavy weather damage, ice damage, hull cracking, or suspected hull defect resulting in immobilization of main engines, extensive accommodation damage, severe structural damage, rendering the ship unfit to proceed, or regardless pollution or a breakdown necessitating towage or shore assistance.
Less serious accidents (LSA)	The marine incident falls outside serious and very serious accidents
Marine incidents	The event which causes no material or moral damage.

According to the IMO, the severities of accidents are based on the three-level categorisation. This categorisation depends on the impact of an event (accident) on people, property and environment (Valdez Banda et al., 2015).

4.2. *Classifiers.* Artificial Intelligence (AI) is a major discipline that includes subdisciplines such as fuzzy logic, artificial neural networks, genetic algorithms, particle swarm optimisation, etc. There are several fields to apply these methods such as classification and regression. In this paper some AI techniques are used for classifying accident data.

Every method has advantages and disadvantages. To reduce the disadvantages, some of these algorithms are merged to achieve new methods. In this paper, we used neuro-fuzzy classifiers (Adaptive Neuro-Fuzzy Classifier (ANFC), Adaptive Neuro-Fuzzy Classifier Using Linguistic Hedges (ANFC-LH)) which are the combinations of fuzzy logic and artificial neural networks, and Genetically Optimised Fuzzy Classifier (GOFC) and compared the yields with Artificial Neural Networks (ANN).

In this study, only AI-based methods are used due to insufficient samples for some accident types. Therefore, the statistical or stochastic classifiers were not used. The ANFC-LH and GOFC methods are initially used to classify sea accidents.

4.2.1. *Adaptive Neuro-Fuzzy Classifier.* Fuzzy Inference Systems (FISs) are the closest method to human thinking. ANNs have learning capabilities. When a problem is modelled with FIS, the FIS parameters can be adapted by ANN training approximation. The ANFC combines FIS and ANN methods. Jang (1993) proposed ANFIS as an approximation function and estimator. ANFIS has only one output. Network-based classifiers should have more than one output. However, many researchers have wrongly used the ANFIS as a classifier. Jang also proposed Adaptive Neuro-Fuzzy Classifiers. Cetisli and Barkana (2010) realised the ANFC using a Scaled Conjugate Gradient (SCG) training algorithm for classification purposes. In neuro-fuzzy classification methods, the feature space is partitioned into multiple fuzzy subspaces that are controlled by fuzzy if-then rules. These rules can be represented by a network structure. However, to determine an optimum fuzzy region, the parameters of the fuzzy rules should be optimised (Jang, 1993; Cetisli and Barkana, 2010). This classifier has both multiple inputs and outputs. The use of weights in the defuzzification step affects the rules and improves the classification flexibility. The K-means clustering method is used to obtain the initial parameters and to formulate the fuzzy if-then rules (Jang et al., 1997; Cetisli, 2010b).

A fuzzy classification rule R_i which describes the relation between the input feature space and the classes can be defined as follows:

$$R_i : \text{IF} x_{s1} \ is \ A_{i1} \text{AND} \ldots x_{sj} \ is \ A_{ij} \text{AND} \ldots x_{sn} \ is \ A_{in} \text{ THEN } class \text{ is } C_k, \quad (1)$$

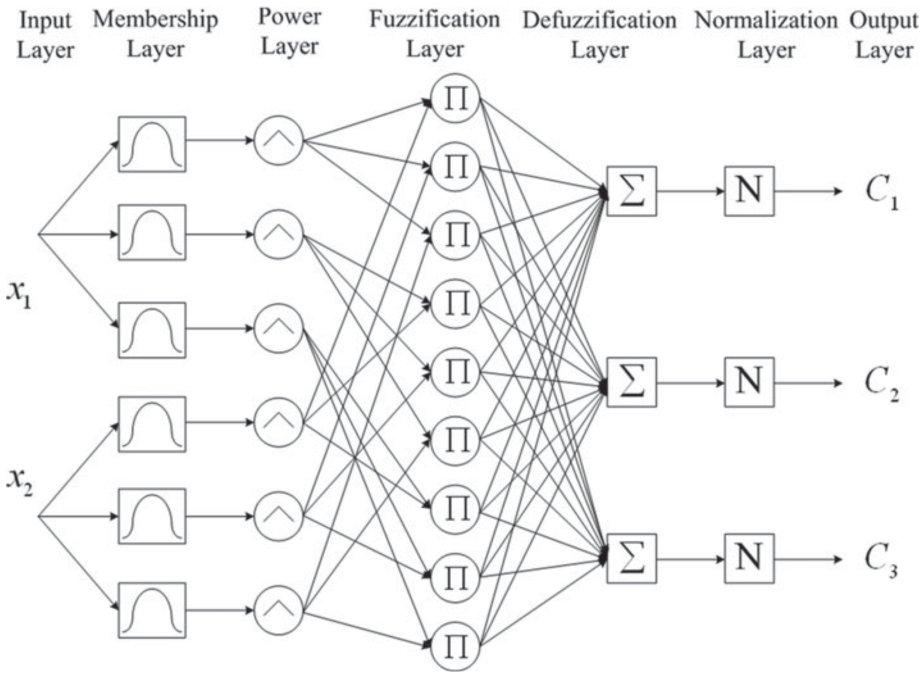

Figure 3. Structure of ANFC-LH Method (Cetisli, 2010a).

where x_{sj} denotes the j^{th} feature or input variable of the s^{th} sample; C_k represents the k^{th} label of class; n represents the number of features; A_{ij} denotes the fuzzy set of the j^{th} feature in the i^{th} rule and is characterised by the appropriate membership function (Jang et al., 1997).

4.2.2. *Adaptive Neuro-Fuzzy Classifier Using Linguistic Hedges.* In the ANFC-LH method, the LHs that are constituted by the power of fuzzy sets introduce the importance of the fuzzy sets for fuzzy rules (Cetisli, 2010b). They can also change the primary meaning of fuzzy membership functions to a secondary meaning. To improve the meaning of fuzzy rules and classification accuracy, a layer, which defines the adaptive linguistic hedges, is added into the proposed classifier network. The LHs are trained with other network parameters by an SCG training algorithm. The tuned LH values of fuzzy sets improve the flexibility of fuzzy sets. This property of LH can improve the distinguishability rates of overlapped classes. Linguistic hedges improve the meaning of linguistic variables and terms.

When the fuzzy rule is redefined with LHs, the new rule occurs as:

$$\begin{aligned} R_i : \text{IF} x_{s1} \ is \ A_{i1} \ \text{with} \ p_{i1} \ \text{AND} \ldots x_{pj} \ is \ A_{ij} \ with \ p_{ij} \ \text{AND} \ldots \\ x_{pn} \ is \ A_{in} \ with \ p_{in} \ \text{THEN} \ class \ \text{is} \ C_k \end{aligned} \tag{2}$$

When the p_{ij} value is equal to 0 value, the A_{ij} fuzzy set is not used. By this way, the rule is shortened. This result can be used for feature selection (Cetisli, 2010a). The structure of the ANFC-LH method is described in Figure 3.

4.2.3. *Feature Selection Using ANFC-LH.* This paper presents a fuzzy Feature Selection (FS) method based on the LH concept. The values of LHs can be used to show the importance degree of fuzzy sets. When this property is used for classification problems, and every class is defined by a fuzzy classification rule, the LHs of every fuzzy set denote the importance degree of input features (Cetisli, 2010a). If the LHs values of features are close to concentration values, these features are more important or relevant, and can be selected. In contrast, if the LH values of features are close to dilation values, these features are not important, and can be eliminated. According to the LHs value of features, the redundant, noisy features can be eliminated, and significant features can be selected. For this aim, a new LH-based FS algorithm is proposed by using ANFC.

4.2.4. *Artificial Neural Networks.* In machine learning and AI, ANN is a popular model inspired by biological neural networks and is used to estimate approximate functions or in classification. Artificial neural networks are generally presented as a system of interconnected neurons which move messages between each other. The connections have numeric weights that can be tuned based on experience, making neural nets adaptive to inputs and capable of teaching (Bishop, 1995).

ANNs are generally made up of three layers: the input, hidden, and output layers. The input layer receives the initial values of the variables, the output layer shows the results of the network for the input, and the hidden layer carries out the operations designed to achieve the output. The number of neurons in the input layer must correspond to the number of input variables, and the output layer must have as many neurons as the number of outputs produced by the network. However, there is no rule to allow prior decisions to be made on the number of neurons the hidden layer should contain or whether the hidden layer must consist of more than one sub-layer. The hidden layer is obtained through trial and error.

Neural computation based on artificial neural networks involves database training to predict the input of the initial values of the variables, while the output layer shows the results of the network for two-dimensional problems, because it is able to imitate the learning capability of human beings. Thus, the network learns directly from the examples without any prior formula regarding the nature of the problem and can independently generalise knowledge, which then can be applied to new cases (Wang et al., 2015). In this paper a feed-forward back propagation neural network is used as the ANN model.

4.2.5. *Genetically Rule Extracted Fuzzy Inference System.* Another approach for FIS generation according to dataset is genetic fuzzy systems, and many methods and algorithms have been reported (Herrera, 2008). In this study, to generate a Sugeno-Typed FIS (Takagi and Sugeno, 1985), a method composed of three stages is followed. In the first stage, fuzzy inputs and outputs are generated according to a training dataset. In the second stage, fuzzy rules are generated according to the dataset via the genetic algorithm approach and in the last stage the numbers of these rules are also reduced via genetic algorithm.

According to this algorithm; cluster size, training dataset and genetic algorithm parameters (chromosome number, crossover and mutation ratios etc.) are received from users. Each feature vector in a dataset is clustered by a fuzzy c-means algorithm. Each cluster centre is signed as a Gauss-typed membership function's centre. Functions are enlarged until the function covers a defined percentage of data. So the first stage of the algorithm is completed and fuzzy inputs and output are generated.

In the second stage the first task is defining the chromosome structure. Every rule is represented as an array of integers. Each item of the array describes the order of membership functions in the fuzzy input or output. As every fuzzy input has membership functions according to cluster number, cluster number is binarized and digit number is achieved and every membership function is represented by bits. With merging binarization rules,

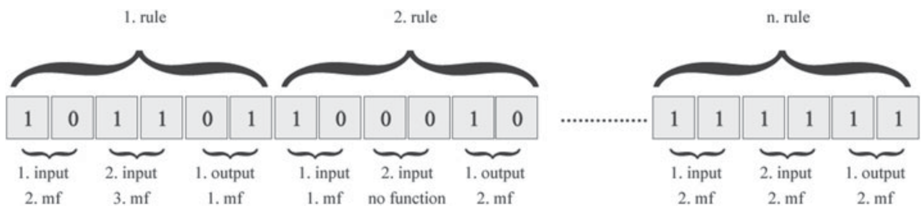

Figure 4. Chromosome Structure of GOFC.

a chromosome is generated. In Figure 4, a sample chromosome structure is shown for FIS which has three inputs, one output, n rules and three clusters. Because cluster number 3 is equal to $(11)_2$ every membership function is represented by two bits.

After the chromosome structure is designed, the fitness function takes an important place in this method. In this paper, mean square error is used as a fitness function. For a fitness result, the chromosome is converted to a rule set and a new rule set is applied for all examples in the training dataset. FIS results are compared with real example values and mean squared error is measured.

The last stage of the algorithm is the rule reduction stage. The chromosome structure of this stage is simpler than this previous stage. Every rule is represented by one bit in the chromosome. If a n. bit takes value 0 this means the n. rule will be removed from the rule set. Like the previous stage, mean squared error is used as the fitness function.

4.2.6. *Give&Take Cross-Validation.* AI-based methods need training to map input-output data. Therefore, the training and testing data are obtained from the original data. If the behaviours of training and testing data are not similar, the adapted model cannot be classified or approximated to the testing data. To avoid this problem, k-fold cross-validation methods are used for the evaluation of models. In k-fold cross-validation, the folds are created randomly. Therefore there is no similarity between training and testing data. Cetisli and Kalkan (2014) proposed a new validation method to satisfy similarity. The Give&Take Cross-Validation (GTCV) is inspired from street games (Cetisli and Kalkan, 2014). In mutual games, the teams should be in equilibrium. So, the team members are selected for neighbourliness values. All data are matched as couples according their neighbourliness values. While one of the couple is put in the training set, the other is put in the testing set. By this way, two similar and balanced groups are obtained.

5. EXPERIMENTAL STUDIES AND DISCUSSION. The existing 135 samples are separated into two datasets, 68 of them as training and 67 of them as testing using the GTCV method. The difference between Random Selection (RS) and GTCV are demonstrated in Figure 5. The first column denotes the training sets, the second column denotes the testing sets. The last column denotes the overlapping of sets in Figure 6. The rows denote GTCV and RS, respectively. Figure 7 shows that the overlapping rate of GTCV is better than RS. This result improves the recognition success. ANFC, ANFC-LH, ANN, and GOFC are adapted using the training dataset. The test data is used for checking the success of the trained methods.

As shown in Figure 7, sometimes, RS can create a smaller training region than the testing area. Therefore, testing accuracy can be decreased by methods that are trained with this

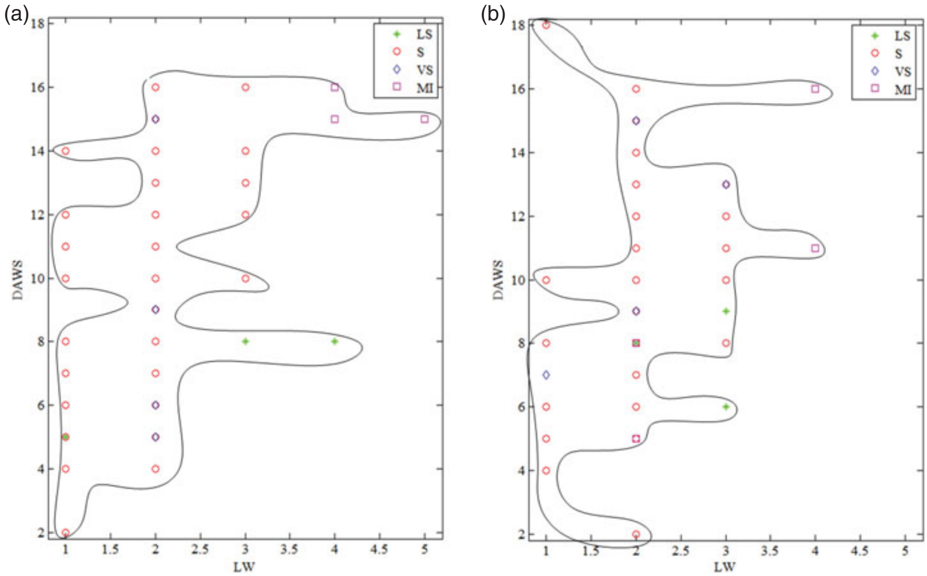

Figure 5. Training and Testing of GTCV (a) Training of GTCV, (b) Testing of GTCV.

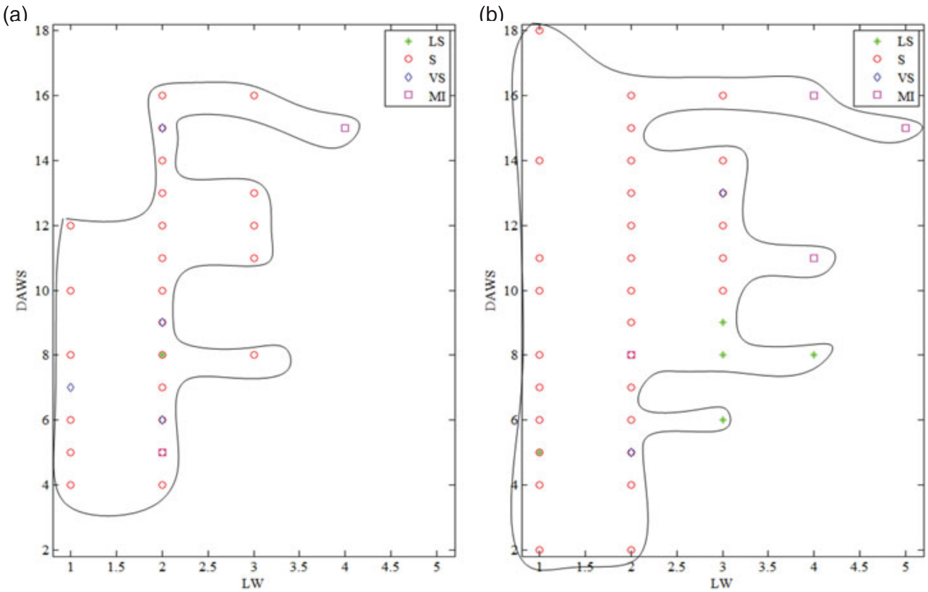

Figure 6. Training and Testing of RS (a) Training of RS, (b) Testing of RS.

training data. This situation is not wanted. The GTCV technique eliminates this situation and it guarantees the high rate of set overlap.

The classification results of methods for testing set of maritime accident data were compared and the results are given in Table 3. The results were obtained with the GTCV method.

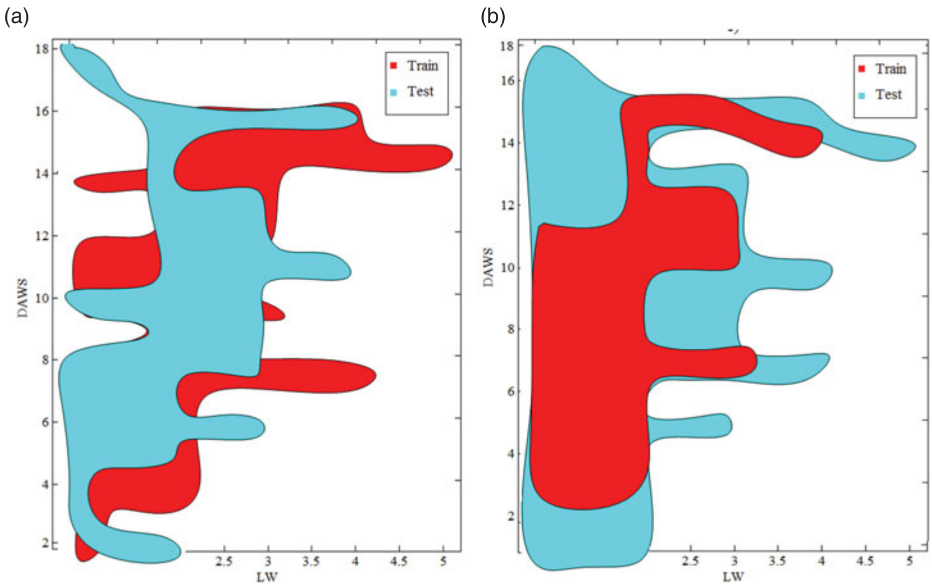

Figure 7. Overlap of GTCV and RS (a) Overlap of GTCV (b) Overlapping of RS.

Table 3. Classification results of maritime accident testing data with GTCV.

Method	Recognition Rate (%)	Explanation
ANFC	89·55	4 rules, 20 neurons
ANFC-LH	89·55	4 rules, 20 neurons
ANN	85·07	4-12-4 network, 20 neurons
GOFC	86·43	7 rules

Table 4. Classification results of maritime accident testing data with RS.

Method	Recognition Rate (%)	Explanation
ANFC	88·05	4 rules, 20 neurons
ANFC-LH	88·05	4 rules, 20 neurons
ANN	83·58	4-12-4 network, 20 neurons
GOFC	84·03	7 rules

According to Table 3, ANFC, ANFC-LH and GOFC give similar results, but ANFC and ANFC-LH used fewer parameters than GOFC. This study was also repeated for RS of the training and testing sets. The experimental results denote the average values of running the classification results ten times. Table 4 represents the classification results of compared methods using random selection.

Table 4 shows that the recognition rates of methods are decreased if RS is used to create training and testing sets. As a result, GTCV achieves a positive contribution to classification results.

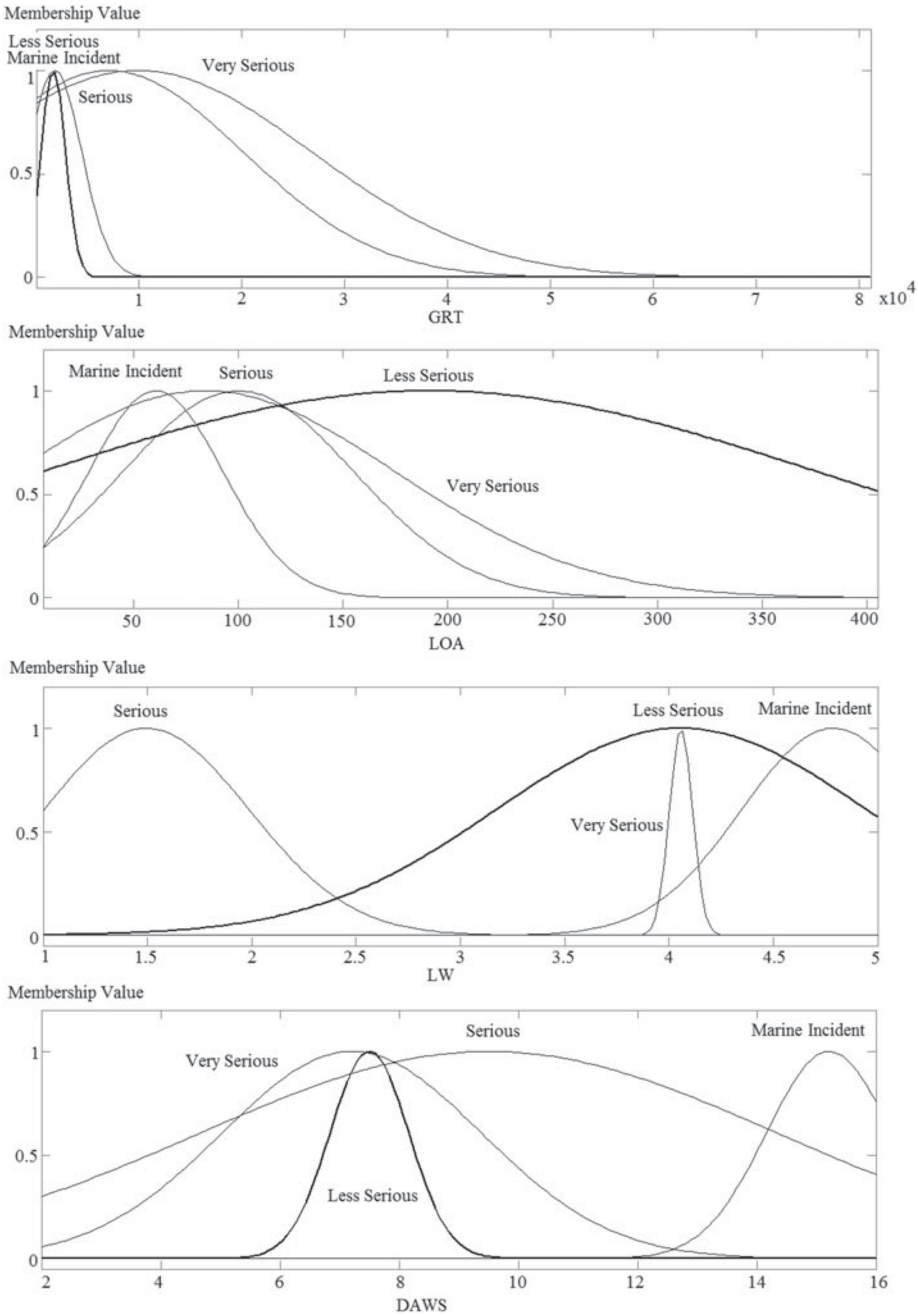

Figure 8. Fuzzy membership functions of inputs.

The fuzzy membership functions for four features and four classes are plotted in Figure 8. There are four rules for four classes and they are also depicted in Figure 9 for any random sample.

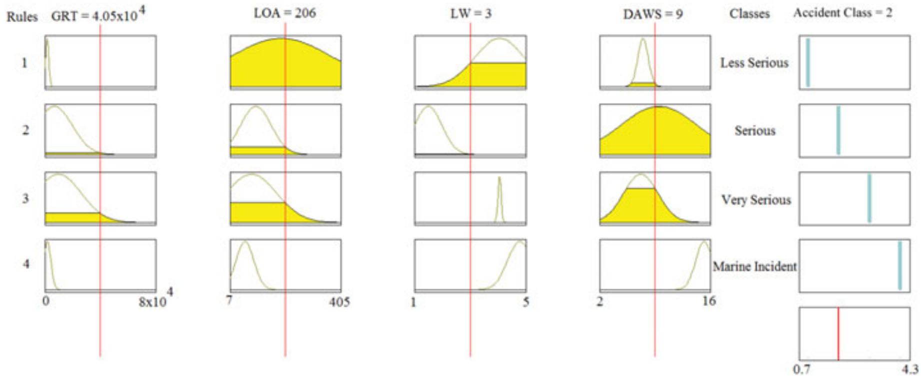

Figure 9. Demonstration of fuzzy rules working for any input sample of ANFC.

Table 5. Linguistic Hedge values of class membership functions.

Features	GRT	LOA	LW	DAWS
Classes				
Less Serious	0·66	1·00	1·00	1·00
Serious	0·00	0·00	1·00	0·40
Very Serious	1·00	1·00	1·00	1·00
Marine Incident	1·00	1·00	1·00	1·00
Total	2·66	3·00	4·00	3·40

Table 6. Best Classification results of maritime accident testing data with LW, LOA and DAWS features.

Method	Best Recognition Rate (%)	Explanation	Used Features
ANFC	85·07	4 rules, 20 neurons	LW, DAWS
ANFC-LH	85·07	4 rules, 20 neurons	LW, DAWS
ANN	86·56	4-12-4 network, 20 neurons	LW, DAWS
GOFC	92·53	7 rules, 6 clusters	LW, DAWS, LOA

If Figure 9 is analysed, the GRT and LOA features conflict for some accident types with overlapping membership functions. In this figure, the fuzzy inference processes is demonstrated. The red vertical lines represent the inputs. The points of interceptions of inputs and membership functions are filled in yellow (fuzzifying inputs). For every four rules, results are shown by vertical grey lines (applying the fuzzy operator). The last process aggregates all outputs and the defuzzifying. The short red line in the bottom right of the figure represents the result of fuzzy inference. This result is also seen with linguistic hedge values of class membership functions in Table 5 below.

Table 5 shows that the gross registered tonnage is not as important in detecting the accident type as other features. The sea wave height and daily average wind speed are more important than others. As a result, the natural conditions affect accidents more in the Istanbul Strait. So with removing the GRT several experiments are made. For every classifier the best results and features are given in Table 6.

Although only two features were used, the obtained results are similar to the previous results for ANFC-LH. The result shows that GRT and LOA are less important than the

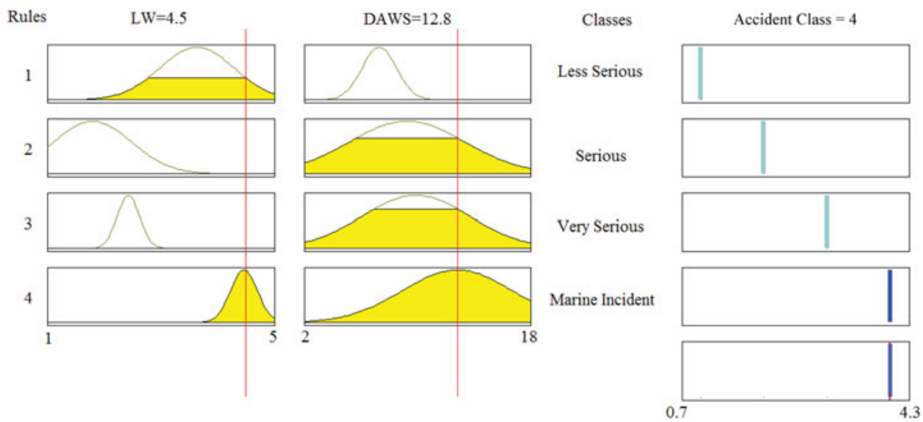

Figure 10. The fuzzy rules of ANFC with two selected features.

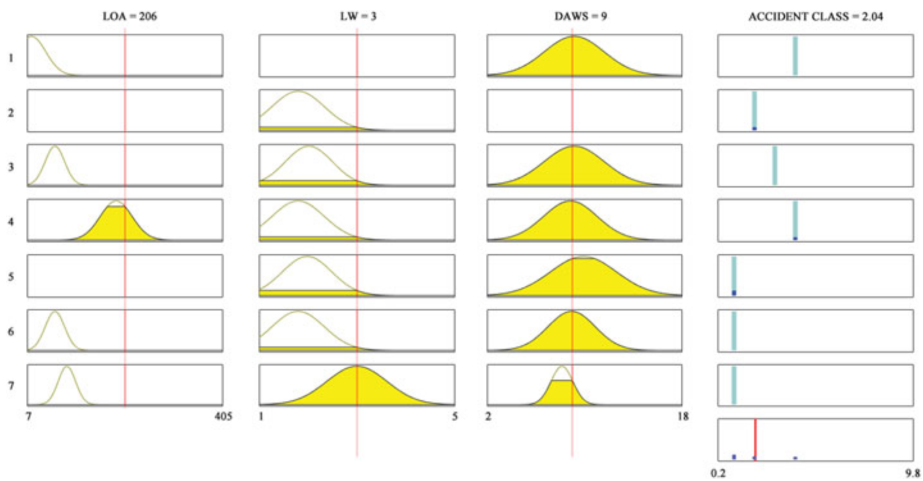

Figure 11. The fuzzy rules of GOFC with three selected features.

others for this method. Figure 10 shows the fuzzy rules of ANFC-LH for the selected features.

The obtained results via ANFC-LH can be abstracted with below equal weighted if-then rules.

- If GRT<6 tonnes and LW>2 m and 2<DAWS<3 m/s then accident type is Less Serious
- If LW<2 m and 2<DAWS<18 m/s then accident type is Serious
- If 2<LW<3 m and 2<DAWS<3 m/s then accident type is Very Serious
- If LW>4 m and DAWS>4 m/s accident type is Marine Incident

On the other hand, the fuzzy rules of GOFC are shown in Figure 11.

When the genetic algorithm is performed, the genes of the membership functions of the GRT parameter are optimised as 0 (zero). So, it is automatically disabled. In addition, equal weighted if-then rules which are generated by GOFC are shown below.

- If LOA is short and DAWS is fast then Accident Type is Very_serious_accidents
- If LW is low then Accident Type is Less_serious_accidents
- If LOA is middle and LW is middle and DAWS is fast then Accident Type is Serious_accidents
- If LOA is very_long and LW is low and DAWS is fast then Accident Type is Very_serious_accidents
- If LW is middle and DAWS is very_fast then Accident Type is Marine_incidents
- If LOA is middle and LW is low and DAWS is fast then Accident Type is Marine_incidents
- If LOA is long and LW is high and DAWS is slow then Accident Type is Marine_incidents

As shown in the if-then rules above, although the LOA is important in determining the ship accident type on narrow waterways, they are not of more value than weather conditions for Istanbul Strait accidents. As seen in the figures above, most of the membership functions are overlapped by others for LOA, and there are no distinctive boundaries between the accident classes. Analysis also indicates that even appropriate parameters such as LW and DAWS can be major factors in serious maritime accidents as shown in the second set of if-then rules.

This situation can be explained by lack of pilotage. Pilots are of utmost importance for safe passage (Ucan and Nas, 2016). Lack of a sufficient and experienced pilotage service significantly increases the accident risks in the Strait (Poyraz, 1998; Akten, 2004; Yurtören, 2004; Ece, 2005; Ulusçu et al., 2009; Küçükosmanoğlu, 2012; Erol and Başar, 2015; Uğurlu et al., 2015). It is clearly seen from research that the human factor is the most important cause of marine accidents (Reason, 1997; Portela, 2005; Çelik and Çebi, 2009; McKnight et al., 2007; Erol and Başar, 2015; Uğurlu et al., 2015). So, Ulusçu et al. (2009) suggested that mandatory pilotage should be implemented for vessels that pass through the Strait and it would be recommended for vessels longer than 150 m or 100 m. They stated that average accident risk would reduce by 7% and by 46% respectively. On the other hand, in this study we have shown that GRT and LOA do not have any distinctive boundaries between the accident classes determined by the IMO. Even though GRT and LOA are not distinctive for these accident classes, increasing ship sizes raise the accidents' impacts on humans, environment and economy (Kristiansen, 2005; Uğurlu et al., 2015). Consequently, in narrow passages such as the Istanbul Strait in which adverse weather conditions can prevail, pilotage services should be mandatory or strongly recommended for all transit vessels regardless of size restrictions. In this context Küçükosmanoğlu (2012) emphasises the necessity of a pilot service through the Istanbul Strait regardless of the size of the vessels. However, according to the Montreux Convention there is no mandatory pilotage for transiting vessels. Thus, aids to navigation in the Strait should be improved consistently and all ships have to follow the passage guidelines.

In their study on the Singapore Strait, Qu et al. (2011) have revealed that speed is a crucial factor in increasing marine accidents. When viewed from this aspect, according to MTRTS (1998), within the Istanbul Strait vessels may not proceed at a speed more than 10 knots over the ground. However, if more speed is needed to maintain good steerage, the nearest traffic control station shall be notified and the master shall proceed with care and caution at a speed which will not create any danger of collision or cause damage by wave making to the banks or properties and other vessels in motion or tied up (MTRTS, 1998. Article 13). However, according to our results, adverse weather conditions have a significant impact on the occurrence of accidents. Therefore, speed and size limitations must be determined in compliance with meteorological conditions for narrow sea passages in which the geographical and meteorological conditions prevail. Thus, the possibility of an unwanted event will be lowered for adverse weather conditions.

6. CONCLUSION AND RECOMMENDATIONS.

In this study, maritime accidents occurring in the Istanbul Strait have been analysed using AI methods such as ANFC, ANFC-LH, ANN and GOFC. Experimental studies show that the ANFC, ANFC-LH and GOFC methods demonstrate similar recognition success. However, ANFC-LH and GOFC rules give us the behaviour of features for different accident types. These rules can be used to develop regulations to assist in reducing their number and severity. The analyses show that there is more than one satisfactory method to reduce accidents.

We have also proposed equal weighted if-then rules. The Ministry of Transport, Maritime and Communications Search and Rescue Department and the Turkish State Meteorological Service databases are used in this study after a data-cleansing procedure. It can be concluded that as LOA increases, the severity of marine accidents rises. Also, as weather conditions deteriorate, the severity of marine accidents can increase. Thus, weather conditions and the ship size should be controlling parameters for the Turkish authorities.

In order to decrease the number and consequences of maritime accidents in the Istanbul Strait, the following precautions that have been developed according to the abovementioned rules may be helpful: (1) Maritime traffic in the Istanbul Strait should be rearranged according to weather conditions. (2) New anchoring areas for vessels could be defined to avoid the effects of bad weather conditions for ships waiting to pass through the Istanbul Strait. (3) Piloting and escorting services should be encouraged for ship masters with limited experience in narrow waters. Indeed, the IMO should be lobbied to make compulsory legal arrangements for taking pilots on narrow and dangerous waterways with intensive ship traffic such as the Istanbul Strait. (4) Safety rules related to weather conditions should be continuously improved. (5) According to if-then rules presented in Section 4, necessary risk reduction solutions such as a ship size limit should be implemented to assure the safe navigation under bad weather conditions in the Istanbul Strait. Due to its strategic importance and traffic density, maritime transit traffic safety in the Istanbul Strait should be increasingly enhanced.

Results obtained from this study can be an important input for risk assessment. The novelty in in this study is to emphasise the importance of weather conditions in the precautions taken to mitigate navigational and environmental risks.

ACKNOWLEDGEMENT

We thank the Ministry of Transport, Maritime Affairs and Communications and Turkish State Meteorological Service for providing us with the data used in this paper.

REFERENCES

Akten, N. (2004). Analysis of shipping casualties in the Bosphorus. *Journal of Navigation*, **57**(3), 345–356.

Arslan, O. and Turan, O. (2009). Analytical investigation of marine casualties at the Strait of Istanbul with SWOT–AHP method. *Maritime Policy & Management*, **36**(2), 131–145.

Aydogdu, Y.V., Yurtoren, C., Kum, S., Park, J.-S. (2010). Questionnaire survey on the risk perception in the Istanbul Strait. *Journal of Navigation and Port Research*, **34**(7), 517–523, http://dx.doi.org/10.5394/KINPR.2010.34.7.517.

Barnett, M.L. (2005). Searching for the root causes of maritime casualties - individual competence or organizational culture? *WMU Journal of Maritime Affairs*, **4**(2), 131–145.

Başar, E. (2010). Investigation into marine traffic and a risky area in the Turkish straits system: Çanakkale Strait, *TRANSPORT*, **25**, 5–10.

Başar, E. and Yıldırım, U. (2014). Effect of the oil areas on marine traffic and oil spill risks at the Black Sea. *Journal of ETA Maritime Science*, **2**(2), 105–110.

Bishop, C.M. (1995). *Neural networks for pattern recognition*. Oxford University Press.

Cetisli, B. (2010a). Development of an adaptive neuro-fuzzy classifier using linguistic hedges: Part 1, *Expert Systems with Applications*, **37**(8), 6093–6101.

Cetisli, B. (2010b). The effect of linguistic hedges on feature selection: Part 2. *Expert Systems with Applications*, **37**(8), 6102–6108.

Cetisli, B. and Barkana, A. (2010). Speeding up the scaled conjugate gradient algorithm and its application in neuro-fuzzy classifier training, Springer. *Soft Computing*, **14**(4), 365–378.

Cetisli, B. and Kalkan, H. (2014). A new deterministic validation method for classifier success: Give&Take, *22nd IEEE Signal Processing and Communications Applications Conference (SIU), Trabzon, Turkey*, 345–348.

Çelik, M. and Çebi, S. (2009). Analytical HFACS for investigating human errors in shipping accidents. *Accident Analysis and Prevention*, **41**(1), 66–75.

Ece, N.J. (2005). *Investigation of accidents in Istanbul Strait, for safety of marine and pollution and determination of passage in safety condition*. Doctorate Thesis, Gazi University.

Erol, S. and Başar, E. (2015). The analysis of ship accident occurred in Turkish search and rescue area by using decision tree. *Maritime Policy & Management*, **42**(4), 377–388.

Goerlandt, F. and Montewka, J. (2015). Maritime transportation risk analysis: Review and analysis in light of some foundational issues. *Reliability Engineering and System Safety*, **138**, 115–134.

Goerlandt, F., Montewka, J., Kuzmin, V. and Kujala, P. (2015). A risk-informed ship collision alert system: framework and application. *Safety Science*, **77**, 182–204.

Harrald, J., Mazzuchi, T., Merrick, J., van Dorp R. and Spahn, J. (1998). Using system simulation to model the impact of human error in a maritime system. *Safety Science*, **30**(1–2), 235–247.

Hashemi, R.R., Le Blanc, L.A., Rucks, C.T. and Shearry, A. (1995). A neural network for transportation safety modelling, *Expert Systems with Applications*, **9**(3), 247–256.

Herrera, F. (2008). Genetic fuzzy systems: taxonomy, current research trends and prospects. *Evolutionary Intelligence*, **1**, 27–46, DOI 10.1007/s12065-007-0001-5.

Jang, J.S.R. (1993). ANFIS: Adaptive network based fuzzy inference systems. *IEEE Transactions on Systems Man and Cybernetics*, **23**(3), 665–685.

Jang, J.S.R., Sun, C.T. and Mizutani, E. (1997). *Neuro-fuzzy and soft computing*. Upper Saddle River: Prentice Hall.

Kristiansen, S. (2005). *Maritime transportation safety management and risk analysis*. Burlington, Biddles.

Kum, S., Furusho, M. and Iwasaki, H. (2008). Investigation on the factors of VTS operators' mental workload: case of Turkish operators. *International Maritime Lecturers Association 16th Conference, Izmir.*

Küçükosmanoğlu, A. (2012). *Maritime accidents forecast model for Bosphorus*. Doctorate Thesis, Middle East Technical University.

Le Blanc, L.A., Hashemi, R.R. and Rucks, C.T. (2001). Pattern development for vessel accidents: a comparison of statistical and neural computing techniques. *Expert Systems with Applications*, **20**, 163–171.

Marine Traffic. (2016). Marine Traffic Istanbul Strait Traffic Density Map. http://www.marinetraffic.com/tr/ais/home/centerx:29/centery:41/zoom:8. Accessed 15 February 2016.

Martins, M.R. and Maturana, M.C. (2010). Human error contribution in collision and grounding of oil tankers. *Risk Analysis*, **30**(4), 674-698, http://dx.doi.org/10.1111/j.1539-6924.2010.01392.x.

McKnight, A.J., Becker., W.W., Pettit, A.J. and McKnight, A.S. (2007). Human error in recreational boating. *Accident Analysis and Prevention*, **39**, 398–405.

Mentes, A., Akyildiz, H., Yetkin, M. and Turkoglu, N. (2015). A FSA based fuzzy DEMATEL approach for risk assessment of cargo ships at coast and open seas of Turkey. *Safety Science*, **79**, 1–10.

Montreux Convention. (1936). Traduction - Translation Convention Regarding The Regime Of The Straits. Signed At Montreux, July 20th 1936.

MTRTS. (1998). Republic of Turkey, Ministry of Transport, Maritime Affairs and Communications, Maritime Traffic Regulations for the Turkish Straits.

Portela, R. de la C. (2005). Maritime casualty analysis as a tool to improve research about human factors on maritime environment. *Journal of Maritime Research*, **2**(2), 3–18.

Poyraz, Ö. (1998). *Coastal crisis management following vessel casualties and application of this in the Turkish Straits*. Doctorate Thesis, Natural and Applied Sciences, Istanbul University.

Qu, X., Meng, Q. and Suyi, L. (2011). Ship collision risk assessment for the Singapore Strait. *Accident Analysis and Prevention*, **43**, 2030–2036.

Reason, J. T. (1997). *Managing the risks of organizational accidents*. Hants: Ashgate.

Takagi, T. and Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, **15**(1), 116–132.

Trucco, P., Cagno, E., Ruggeri, F. and Grande, O. (2007). A Bayesian belief network modelling of organizational factor in risk analysis: A case study in maritime transportation. *Reliability Engineering and System Safety*, **93**(6), 823–834.

Ucan, E. and Nas, S. (2016). Analysing Istanbul Strait Maritime Pilot Capacity by Simulation Technique. *The Journal of Navigation*, **69**, 815–827. http://dx.doi:10.1017/S0373463315000909

Uğurlu, Ö., Erol, S. and Başar, E. (2015). The analysis of life safety and economic loss in marine accidents occurring in the Turkish Straits. *Maritime Policy & Management*, **43**(3), 356–370. http://dx.doi.org/10.1080/03088839.2014.1000992.

Uğurlu, Ö., Yıldırım U. and Yüksekyıldız, E. (2013). Marine Accident Analysis with GIS. *Journal of Shipping and Ocean Engineering*, **3**(12), 21–29.

Ulusçu, Ö.S., Özbaş, B., Altıok, T. and Or, İ. (2008). *Risk analysis of transit vessel traffic in the strait of Istanbul.* Modelling Technology Policy, Laboratory for Port Security.

Ulusçu, Ö.S., Özbaş, B., Altıok, T. and Or, İ. (2009). Risk analysis of the vessel traffic in the strait of Istanbul. *Risk Analysis*, **29**(10), 1454–1472. http://dx.doi.org/10.1111/j.1539-6924.2009.01287.x

Valdez Banda, O.A., Goerlandt, F., Montewka, J. and Kujala, P. (2015). A risk analysis of winter navigation in Finnish sea areas. *Accident Analysis and Prevention*, **79**, 100–116.

Wang, J., Pillay, A., Kwon, Y.S., Wall, A.D. and Loughran, C.G. (2005). An analysis of fishing vessel accidents. *Accident Analysis and Prevention*, **37**(6), 1019–1024.

Wang L. Zeng Y. and Chen, T. (2015). Back propagation neural network with adaptive differential evolution algorithm for time series forecasting, Expert Systems with Applications, 42(2), pp. 855–863.

Wei, L., Hu, Z., Dong, L. and Zhao, W. (2015). A damage assessment model of oil spill accident combining historical data and satellite remote sensing information: A case study in Penglai 19-3 oil spill accident of China, Marine Pollution Bulletin, 91, pp. 258–271.

Yazici, M.A. and Otay E.N. (2009). A Navigation Safety Support Model for the Strait of Istanbul, The Journal of Navigation, 62, pp. 609–630, http://doi:10.1017/S0373463309990130

Yıldırım U., Uğurlu, Ö. and Başar, E. (2015). Human error in grounding accidents: case study for container ships, Journal of ETA Maritime Science, **3**(1), pp. 1–10, http://dx.doi.org/10.5505/jems.2015.57966

Yurtören, C. (2004). Marine traffic management in Istanbul Strait, Doctorate Thesis, Kobe University.